

論文の内容の要旨

論文題目	最大クリーク問題を用いた 複数等質テスト自動構成手法とその近似手法
学位 申請者	石井 隆稔

本論文ではeテストイングにおける複数等質テスト自動構成手法を提案する。

一般に、資格試験や入学試験など、その結果が受験者に重要な影響を与えるテストをハイステークス・テストと呼ぶ。このハイステークス・テストでは、複数等質テストが必要になる場合が多い。複数等質テストとは、異なる項目により構成されているにもかかわらず、各テストが等質であるようなテスト集合である。例えば、毎年行われる資格試験において、毎回のテストで合格難易度が異ならないよう、得点分布および回答所要時間分布などのテストの特性がほぼ一定でなければならない。既に現在、e テスティングの実施に関する標準規格ISO/IEC 23988で、ハイステークス・テストでの複数等質テスト構成が条件として記載されており、複数等質テストの重要性は日々増している。

eテストイングにおけるテスト構成では、アイテムバンクと呼ばれる項目（出題可能な問題）データベースを用いる。このデータベース中には項目内容・項目ごとの出題領域や統計データなどが格納されている。一般に新規項目の作成はeテストイングシステム構成作業で最もコストが高く、一項目を実用化するためにもデータ収集を含めて数ヶ月を要する場合もある。その為、複数等質テスト構成の研究において重要な課題の一つは、アイテムバンクを有効利用するため与えられたアイテムバンクからなるべく多くのテストを構成することである。

本研究では、与えられたアイテムバンク中からできる限り多くのテストを構成可能な複数等質テスト自動構成手法を提案する。

まず、厳密に最大数のテストを構成する手法を提案する(以降厳密法と呼ぶ)。具体的には、複数等質テスト構成を最大クリーク問題として定式化する。最大クリーク問題とはグラフ理論の最適化問題であり、与えられたグラフ中からクリーク(どの二つの頂点も接続されている構造)を探索する問題である。本手法ではテストを頂点とみなし、等質なテスト間を接続したグラフから最大クリークを探索することで、最大の複数等質テスト群を出力可能である。この様に構成されたグラフ中のクリークは複数等質テストである。なぜならば、このクリーク中では任意の2頂点は接続されており等質である。従って、このグラフ中の要素数最大のクリークは最大の複数等質テスト群となる。

ただし、厳密法の計算量は与えられたアイテムバンクの大きさに対し超指数的に増大し、大規模なテスト構成は困難である。なぜならばテストは項目の組み合わせであるため、出題可能な項目数が増えればテスト構成条件を満たす可能テスト数は組み合わせ爆発的に増える。この可能テストは等質テスト群を表すグラフの頂点であるため、このグラフ頂点数も組み合わせ爆発的に増え、保持に必要な空間計算量が非常に大きくなる問題がある。

そこで、厳密法の一部を、乱数探索を用いて近似化し計算量の問題を緩和した近似手法を提案する。具体的には、等質性を満たすグラフ全域からの最大クリーク探索を行うのではなく、無作為に抽出した部分グラフからの最大クリーク探索を繰り返すことで、漸近的にグラフ全域の最大クリークを探索する。具体的には、まず無作為に部分グラフをグラフの全域から抽出する。そしてこの部分グラフ内の最大クリークを探索し、現在見つかっている最も大きなクリークと比較、より大きな方を保持する。この探索と現在見つかっている最大の更新を繰り返すことで、本手法は、一度に探索される部分グラフの大きさが全域での最大クリークよりも大きな条件で、時間漸的に全域での最大クリークを発見可能である。このアルゴリズムにより、厳密法の時間計算量 $O(2^{0.1917F})$ 、空間計算量 $O(F^2)$ をどちらも $O(1)$ へと軽減した。ただし、 F は条件を満たすテスト数である。このような近似により、本近似手法はより少ない計算コストで厳密法と同等のテスト数を構成可能である。

本論文ではこれらの手法の有効性をシミュレーションデータおよび実データを用いた実験により示した。具体的には、様々な条件で従来手法とのテスト構成数比較を行った。比較を行った手法は、現在最も普及しているLinden(2005)、提案近似手法と同じく乱数探索を用いた手法であるSun et al. (2008)、日本の情報処理技術者試験に用いられているSongmuang and Ueno(2011)である。多くの場合において、本手法は従来手法以上のテスト数を構成可能であることを示した。また、アイテムバンクサイズが多くなる場合、どの手法でもテスト構成数は増加するが、その増加数は提案手法が最も多い結果となった。つまり、大規模なテスト構成において、提案手法は最も効率よくテスト構成数を増やすことが可能であり、従来手法を大幅に改善することを示した。

論文審査の結果の要旨

学位申請者氏名 石井 隆稔

審査委員主査 植野 真臣

委員 大須賀 昭彦

委員 栗原 聡

委員 田原 康之

委員 古賀 久志

石井氏の学位論文では、e テスティングにおける複数等質テスト自動構成手法を提案・開発している。

第1章では資格試験などのハイステークス・テストをe テスティングで行う場合、なるべく大量の複数等質テストが必要であることをまとめている。

第2章では、複数等質テスト構成の先行研究を紹介している。実用されている3つの手法と本論文と関係深い1つの先行研究を紹介している。本章では、テスト構成の先行研究はどれも最適化問題を利用しており、そのモデル化の解決方法が異なることを指摘している。次に実用されている手法は、どの手法も与えられた条件で最大数のテストを構成できないという問題を指摘している。最後に、実用化はされていないが可能な限り多くのテストを構成するBelov and Armstrongの手法についても紹介を行っている。この手法は、テスト間に項目の重複を許さないという条件でのみ最適化問題の利用が可能であり、この条件がテスト構成数を著しく制限するため実用的には用いられていないことを指摘している。実用されている手法はテスト間の項目重複を許すことで、テスト構成数を数倍から数十倍にまで増やすことが可能であり、実用化のためにはテスト間に項目の重複を許した条件でのテスト構成が必須であることをまとめている。

第3章ではテスト間に重複を許す条件において厳密に最大数の複数等質テストを構成する手法を提案している。本手法は第2章で紹介したBelov and Armstrongの手法をの重複項目条件について一般化したものであると述べている。

本手法はテスト構成を最大クリーク問題として行う手法である。最大クリーク問題は Belov and Armstrongの手法で使用されている集合充填問題の一般化である。具体的には、テストを頂点とし等質で重複項目数が一定数以下なテスト間に辺を引いたグラフからの最大クリーク問題として複数等質テスト構成をおこなう。この定式化により、与えられたアイテムバンク・テスト構成条件から厳密に最大数の複数等質テストを出力する。これらを示すため、実データとシミュレーションデータを用いた実験を行い、従来手法に比べ多くのテストを構成できる有効性を示している。

第3章で提案・開発した厳密手法の最大の問題は、最大クリーク探索を行う際の空間計算量であると述べている。これを解決するため第4章では厳密手法を乱数探索により近似し、計算量の問題を緩和する手法を提案・開発している。厳密手法の問題点は、複数等質テスト構成のためのグラフ構造が非常に大きいため、計算機の主記憶上に保持できない問題であると述べている。そこで本近似手法では、グラフ全域からの探索の代わりに、ランダムに選び出した部分グラフからの探索を繰り返すことにより空間計算量を軽減している。この近似により、グラフ全域をメモリ上に保持する代わりに部分グラフをメモリ上に保持することで、空間計算量を軽減している。また、本手法は時間漸近的に最大クリークを発見可能であることも主張している。最後に本手法の有効性を示すため、シミュレーションデータおよび実データを用いた実験により、テスト構成数の比較を行いその有効性を示している。ただし、テスト構成の規模に対し与える空間計算量条件が少ない場合、従来手法よりもテスト構成数が少なくなる場合があることも指摘している。そのような場合においても、重複項目数を増やすことで従来手法よりも効率的にテスト構成数を増やせることも示している。

最後に、第5章において本論文で得られた主な研究成果を総括し、本論文をまとめるとともに本論文の課題について述べている。

本論では、eテストイングにおいて最も多くのテストを構成できる手法として実用可能な手法を提案している。この手法はリクルート社との共同研究でもSPIなどのテスト構成において実践的に使用されており、その実用性は高い。

よって、本論文は博士（工学）の学位請求論文として十分な価値を有するものと認める。